基于选矿生产数据的矿石可磨性智能识别算法

朱佼佼，李宗平，师本敬，吴革雄，文 武
（中冶长天国际工程有限责任公司，湖南 长沙 410205）

摘要：本文以在线识别矿石可磨度为研究目标，通过详细分析磨矿过程参数间的相互影响关系，得出磨矿过程运行指标的变化能从一定程度上反应矿石可磨度变化的结论。针对工业数据维度高、耦合性强等问题，在着重研究了 KPLSR 和 SAE 两种可提取关键特征算法机理的基础上，以实际采集的磨矿过程数据为样本，对多算法进行横纵对比，最终验证了深度学习算法 SAE 在工业复杂数据分析中的优势。并以测试试验中性能最佳的算法结构作为矿石可磨性智能识别系统的在线识别模型，该模型在工业化应用中绝对值误差平均可达 0.61 kW·h/t，平均误差率约为 4.79%，可满足现场生产的需求。

关键词：矿石；可磨度；磨矿过程，数据驱动，软测量技术，深度学习算法
中图分类号：TD921；TP29 **文献标识码**：A **文章编号**：1000－8764(2024)04－0083－10
doi：10.13403/j.sjqt.2024.04.062

Intelligent identification algorithm of ores grindability based on mineral processing production data

ZHU Jiaojiao，LI Zongping，SHI Benjing，WU Gexiong，WEN Wu
(Zhongye Changtian International Engineering Co.，Ltd.，Changsha 410205，Hunan，China)

Abstract：The research goal is to identify the ore grindability online in this paper. Through a detailed analysis of the interaction between the grinding process parameters，it is concluded that the change of the grinding process operation index can reflect the change of ore grindability to a certain extent. In order to solve the problems of high dimensionality and strong coupling of industrial data，on the basis of focusing on the mechanism of KPLSR and SAE two algorithms that can extract key features，the actual collected grinding process data is used as a sample to compare multiple algorithms horizontally and vertically，and finally the advantages of deep learning algorithm SAE in industrial complex data analysis are verified. The algorithm structure with the best performance in the test is used as the online recognition model of the ore grindability intelligent identification system，and the average absolute error of the model can reach 0.61 kW·h/t in industrial applications，and the average error rate is about 4.79%，which can meet the needs of on-site production.

Key words：Ore；grindability；grinding process；data-driven；soft measurement technology；deep learning algorithm

矿石可磨度是指矿石在指定磨矿条件下被磨碎的难易程度，是矿石重要性质之一。由于采矿场地不同，以及围岩的多少、没有“均一化”堆取等原因，入磨矿石可磨性经常发生变化。矿石可磨性的波动会直接影响磨矿过程质量指标、运行指标和能耗指标。例如，当矿石可磨性向上波动时，矿浆粒度会变粗，磨机负荷和能耗会变大，如果不及时调整磨矿参数，甚至会出现磨机“胀肚”等现象，导致生产的停滞。

目前，矿石的可磨性指标基本是采用 Bond 功指数。然而，Bond 粉磨功指数测定是用若干个循环的间歇式粉磨筛分来模拟连续闭路的粉磨过程，

收稿日期：2023－10－13；**修回日期**：2023－11－23
基金项目：五矿科技专项计划项目(2020ZXC03)
作者简介：朱佼佼(1986—)，女，博士，从事智能算法在冶金工业领域的创新研究和应用。

需要专门的测定设备和一定数量的有代表性的矿样，其操作过程存在一定程度的复杂性，有一定的劳动强度，并需要一周左右的测定时间[1-2]。因此，多年来一些粉碎工作者探索了对其试验进行简化或计算机模拟的方法。例如：推测法，这类方法适用于已知一种物料或一个测定粒度的Bond粉磨功指数，推测其他物料或测定粒度的Bond粉磨功指数[3-5]；部分测定和模拟算法，这一类方法先采用标准Bond功指数磨机和测定方法，进行2~4个循环的测定，然后根据获得数据，用计算机模拟以后各次循环的测定过程，该类方法主要包括Kapur法[6]和赖复兴法[7]；模拟计算方法，这一类方法只需要少量的有代表性矿样或钻探岩芯样品，在实验室磨机中进行一定的粉磨动力学试验，获得被测矿样的模型参数，然后根据一定的粉磨动力学模型，用计算机模拟Bond粉磨功指数测定和数据处理过程，与前面所述的模拟方法之区别是，这一类方法自第一个循环起就采用模拟，是更全面的模拟，例如熊维平等[8]提出的方法可模拟测定过程和物料的破裂过程，获得最终产品粒度分布和80%通过的粒度大小，以此计算Bond功指数；陈炳辰等[9]、仲崇波等[10]提出的模拟方法需要基于回归分析获得计算功指数经验公式；AKSANI等[11]提出的模拟方法的特点是用累积基础的动力学模型模拟Bond球磨功指数，测定物料颗粒的碎裂过程，模型参数需要通过分批粉磨试验确定。这些改进方法虽然在一定程度上减少了试验工作量，缩短了Bond功指数的测定时间，但依然无法满足磨矿流程控制实时性的需求。

现场经验丰富的专家虽然可以根据磨矿流程中其他参数的变化对矿石可磨性做出一些模糊的判断。但磨矿过程复杂多变，矿石可磨性波动无一定规律可循，各过程变量间耦合性、非线性强，人工凭经验判断易受各种因素的影响，且误判率高。数据驱动的软测量技术不受人为主观因素的影响，且具有实时性，已被广泛用于流程工业[12-13]。磨矿流程和其他流程工业一样，采集的数据具有高维度、强耦合、高冗余，即数据量大、质量差、信息量少等特点。目前处理这类数据的方法主要分为两类：一类是结合多元统计学分析法的机器学习算法。多元统计学分析法用于关键特征提取，机器学习算法实现数据非线性拟合。常用的多元统计学特征提取算法主要有主成分分析法(principle component regression，PCR)、基于核函数的主成分分析法(kernel principle component regression，KPCA)、偏最小二乘回归(partial least squares regression，PLSR)、基于核函数的偏最小二乘回归(kernel partial least squares regression，KPLSR)等，KPLSR首先将原始变量空间进行非线性函数投影到高维空间，然后在高维空间中，原始变量的非线性关系转化为高维空间变量的线性关系，从而通过在高维空间采用线性PLSR进行特征提取。该方法虽然解决了PCR和PLSR无法处理非线性问题的缺点，但当样本量过大时，会出现“维数灾难”问题[14-15]。另一类是直接采用深度学习算法。相对于多元统计学算法，深度学习更适合处理数据量大问题。深度学习作为人工智能领域近些年来发展最快的学科，其在众多复杂问题上表现出超强分析能力。目前常用的深度学习算法主要包括堆栈自编码(stacked auto-encoders，SAE)、深度置信网络(deep belief networks，DBN)和深度卷积神经网络(convolutional neural networks，CNN)[16-19]，其主流应用是模式识别问题，例如文字处理、图像语音识别等。但在工业领域，特别是在过程监测和软测量方面，鲜少有相关的研究和应用。

本文在详细分析基于核函数的偏最小二乘算法(KPLSR)和对战自编码算法(SAE)两种可提取关键特征算法机理的基础上，基于采集的磨矿数据对KPLSR、基于高斯函数的径向基神经网络(gaussian radial basis function neural network，Gaussian_RBFNN)和SAE算法进行了多组训练测试，最终以性能最好的模型作为矿石可磨度的预测模型，以实现矿石可磨性的在线识别，及时反馈生产控制。

1 矿石磨矿过程分析

矿石磨矿过程承接着碎矿和浮选两个工序，是碎矿工序的延续，是浮选工序前的最后一道工

序。磨矿的目的是使矿石中有用矿粒与脉石部分或全部分离达到单体状态的同时，又要防止出现“过磨”和“欠磨”现象，达到浮选需要的颗粒大小的要求。磨矿过程一般采用由球磨机与分级设备形成闭路流程，图 1 展示了典型的一段闭路磨矿分级流程，设备主要包括球磨机、旋流器和砂泵池等。

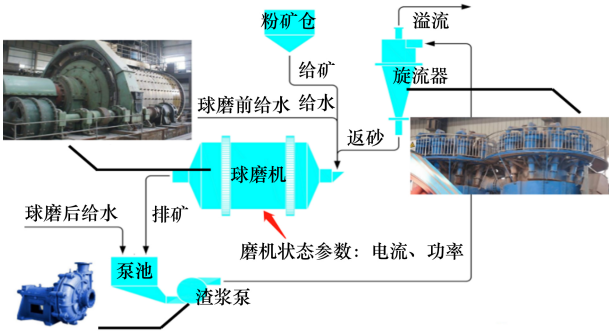


图 1　典型矿石磨矿分级流程

Fig. 1　Grading process of typical ore grinding

生产中，矿石磨矿分级过程的主要影响因素包括：矿石自身的性质、入磨给料粒度、磨机尺寸规格、磨机转速率、钢球密度、尺寸和配比、装球量和补球制度、下料量、料球比和装料量、磨矿浓度和黏度等，这些影响因素之间具有很强的耦合性。例如，磨机入口处的原矿下料量、给水量、返砂量中任何一个量的改变都会引起磨矿浓度的改变，因而又影响分级入口加水量，反过来引起返砂量和返砂浓度的变化。图 2 展示了这些因素之间的相互关系，其中矿石自身性质尤其是可磨性直接影响磨矿分级效率，是影响磨矿作业的关键因素。

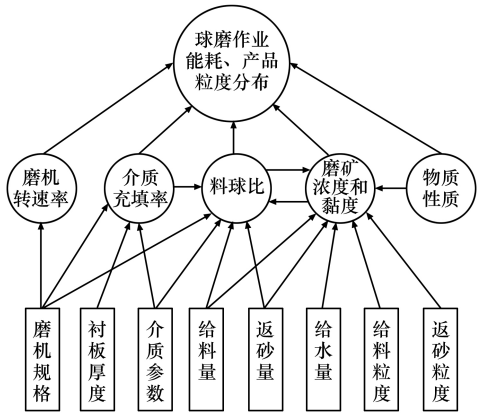


图 2　矿石磨矿分级工艺影响因素关系

Fig. 2　Relationship between the influencing factors of ore grinding and grading process

2　数据驱动的矿石可磨性智能识别方法

矿石可磨度与该流程的设备参数、操作参数和矿石参数之间的关系错综复杂。为了提高分析准确性，需要采集磨矿过程中尽可能多的参数类型。根据专家经验，操作参数主要采集给矿量、给水量、介质量（加球量）、返砂流量和粒度；设备参数主要采集磨机的转速和功率；磨矿产品参数主要采集排矿粒度；另外，矿石性质如粒度和品位也是关键采集数据。图 3 展示了基于数据驱动的矿石可磨性智能识别方法的结构。

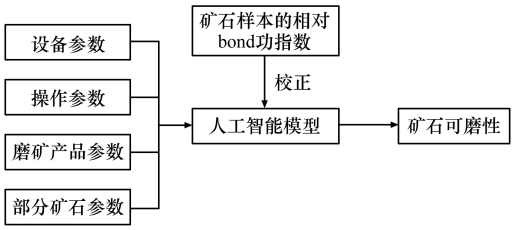


图 3　数据驱动的矿石可磨性智能识别方法结构示意

Fig. 3　Structure of data-driven intelligent identification method for ore grindability

3　智能算法原理分析

3.1　基于核函数的偏最小二乘算法（KPLSR）

核方法具有较好的模型泛化能力，能够有效克服欠拟合和过拟合问题，其将原始变量空间进行非线性函数投影到高维空间，在高维空间中，原始变量的非线性关系转化为高维空间变量的线性关系，从而通过在高维空间采用线性 PCA 和 PLSR 等方法进行特征提取。

目前最常用的核函数是高斯核函数，其表达式如下：

$$k(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{1}$$

式中：σ 参数决定了高斯核函数的宽度。

样本集 $\boldsymbol{X}_{l\times n}$ 的核函数映射计算如下：

$$\boldsymbol{K}(\boldsymbol{X},\boldsymbol{X})=\begin{pmatrix} k(x_1,x_1) & k(x_1,x_2) & \cdots & k(x_1,x_l) \\ k(x_2,x_1) & k(x_2,x_2) & \cdots & k(x_2,x_l) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_l,x_1) & k(x_l,x_2) & \cdots & k(x_l,x_l) \end{pmatrix} \tag{2}$$

式中：$x_i\in\boldsymbol{X}_{l\times n}$，$i=1,2,3,\cdots l$。

样本集被映射到高维空间后，采用 PLSR 提取关键特征，具体步骤：

（1）对映射后的高维数据集 $\boldsymbol{X}_{l\times l}$和目标数据集 $\boldsymbol{Y}_{l\times 1}$进行中心化或标准化处理。

（2）求主成分对。设 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的第一个主成分轴向量分别为 $\boldsymbol{\omega}_1$ 与 $\boldsymbol{c}_1$，由该轴计算得到的 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的第一个主成分：

$$\begin{aligned}\boldsymbol{t}_1 &= \boldsymbol{X}\times\boldsymbol{\omega}_1\\ \boldsymbol{u}_1 &= \boldsymbol{Y}\times\boldsymbol{c}_1\end{aligned}\tag{3}$$

为了 $\boldsymbol{t}_1$ 与 $\boldsymbol{u}_1$ 之间的相关性最大，设目标函数为

$$\max[\rho(\boldsymbol{t}_1,\boldsymbol{u}_1)=\frac{\operatorname{cov}(\boldsymbol{t}_1,\boldsymbol{u}_1)}{\sigma_{t_1}\sigma_{u_1}}]\tag{4}$$

由此可得，$\boldsymbol{t}_1$ 和 $\boldsymbol{u}_1$ 之间协方差越大，各自的方差越小，可满足该目标函数。

然后根据拉格朗日乘子的方法可以求出 $\boldsymbol{\omega}_1$ 和 $\boldsymbol{c}_1$，最后可解的，$\boldsymbol{\omega}_1$ 是对称矩阵 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}$ 的最大特征值对应的特征向量，$\boldsymbol{c}_1$ 是 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$ 的最大特征值对应的特征向量。

（3）求解矩阵 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}$ 和 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$ 和的特征值及特征向量

令 $\boldsymbol{R}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}$ （5）

式中：λ_1，λ_2，…，λ_b 为 $\boldsymbol{R}$ 的 b 个特征根，对方程组 $\boldsymbol{R}\times\boldsymbol{p}=\lambda_p$ 进行求解，得到 $\boldsymbol{P}_1$，$\boldsymbol{P}_2$，…，$\boldsymbol{P}_b$ 为 λ_1，λ_2，…，λ_b 对应的特征向量。

令 $\boldsymbol{Q}=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$，

$$|\boldsymbol{Q}-\boldsymbol{\lambda}\boldsymbol{I}_d|=0\tag{6}$$

式中：λ_1，λ_2，…，λ_d 为 $\boldsymbol{Q}$ 的 d 个特征根，对方程组 $\boldsymbol{Q}\times s=\lambda_p$ 进行求解，得到 $\boldsymbol{S}_1$，$\boldsymbol{S}_2$，…，$\boldsymbol{S}_d$ 为 λ_1，λ_2，…，λ_d 对应的特征向量。

（4）主成分的计算

$$\begin{aligned}\boldsymbol{T}&=\boldsymbol{X}_i\times\boldsymbol{P}_j\\ \boldsymbol{U}&=\boldsymbol{Y}_i\times\boldsymbol{S}_k\end{aligned}\tag{7}$$

（5）主成分贡献率的计算。

CPV 法是基于特征值获取各主成分的方差贡献率，用来体现主成分表达原数据的能力，越大越好。

$$S_{\mathrm{CPV}_i}=\frac{\lambda_i}{\sum_{i=1}^{n}\lambda_i}\tag{8}$$

式中：S_{CPV_i}是第 i 个主成分的方差贡献率；λ_i 是第 i 个主成分所对应的特征值。

（6）求第一主成分对之间的线性拟合模型，判断偏差是否满足拟合需求，不满足，则继续提取第二主成分对。

3.2 堆栈自编码算法(SAE)

堆栈自编码(SAE)是多个自编码模型多级串联起来的结构，每一个自编码模型又包含编码器和解码器两个结构，通过这个结构模型可以提取高维度强耦合工业过程数据的高阶特征，最后利用归回模型拟合提取的关键特征与目标参数之间的关系。其结构如图 4 所示。

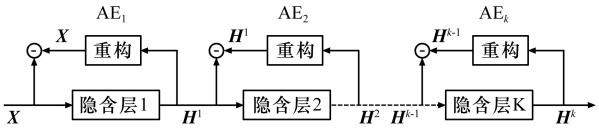
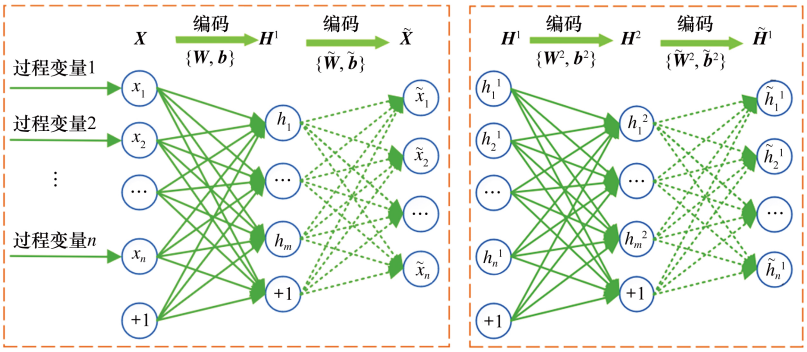


图 4 SAE 深度学习模型结构

Fig. 4 Deep learning model structure of SAE

图 4 中，AE_1 和 AE_2 的具体结构如图 5 所示。



（a）AE_1 的具体结构；（b）AE_2-AE_k 的具体结构

图 5 每一层 AE 的具体结构

Fig. 5 Specific structure of each AE layer

AE_1 的输入为多个工业过程数据 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)$，$AE_1$ 的隐层 $\boldsymbol{H}$ 中每一个节点 $h=f(\omega x+b)$，经过每一层 AE 都会提取出相应的特征 $\boldsymbol{H}^1=(h_1^1,h_2^1,\cdots,h_n^1)$，隐层 $\boldsymbol{H}$ 是对输入数据的特征提取数据，模型的学习目标是使输入向量 $\boldsymbol{X}$ 与输出向量 $\widehat{\boldsymbol{X}}$ 的均方差最小，其原理就是让提出的特征经过解码能尽可能的还原数据，其目标函数如下。

$$T=\min\left(\sum_{i=1}^{n}\parallel\widetilde{x_i}-x_i\parallel^2/2n\right) \tag{9}$$

式中：n 为输入变量的个数。

当多个自编码模型多级串联起来构成堆栈自编码(SAE)，网络的训练分为两个步骤：无监督预训练和有监督的调整过程。

在无监督训练过程中，每一层 AE 顺序单独训练，当 AE_1 训练完，获得初始的$\{\boldsymbol{W}^1,\boldsymbol{b}^1\}$，$AE_2$ 的输入为 AE_1 隐层的输出 $\boldsymbol{H}^1=(h_1^1,h_2^1,\cdots,h_n^1)$，$AE_2$ 的学习过程是以 $\boldsymbol{H}^1$ 为输入和输出的自学习过程，同时提取出新的特征 $\boldsymbol{H}^2=(h_1^2,h_2^2,\cdots,h_m^2)$，获得初始的$\{\boldsymbol{W}^2,\boldsymbol{b}^2\}$，以此类推直到 AE_k 层，输出原数据的高层次特征，且获得网络所有的初始参数。

有监督反向学习过程，在 SAE 的最后堆叠一个无隐层的双层 FFNN 实现与目标参数的拟合，整个网络的训练类似多隐层 BP 神经网络，AE 层的网络初始参数采用预训练获取的网络参数，双层 FFNN 的网络参数采用随机值，以网络输出与目标参数的偏差反向调整网络的连接权值，直到满足训练条件。如图 6 所示，其中，$\boldsymbol{H}^k=(h_1^k,h_2^k,\cdots,h_m^k)$为第 k 个 AE 层提取的特征数据。

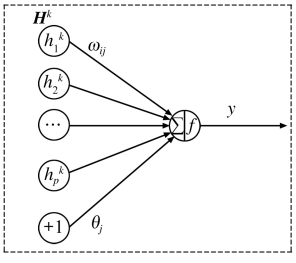


图 6　SAE 深度学习算法拟合层

Fig. 6　Fitting layer SAE deep learning algorithm

4　智能算法测试与对比

4.1　数据集预处理与分析

磨矿过程的数据采集，除了从 DCS 系统采集的生产数据，例如给矿量、给水量、磨机功率等，还包括统计数据：处理量、钢球添加量；检化验数据：溢流粒度、相对可磨度等，这些数据采样周期不统一、检化验数据与生产数据时滞不确定，采集数据因各种原因存在空值、负值、零值等问题，需要经过预处理技术，以确保准确、高效或有意义的分析。对于数据采样周期不统一的问题，这里以最大周期的标准，将所有的数据处理为该采样周期；对于检化验数据与生产数据时滞不确定问题，主要按照现场经验拉齐时间轴，例如取溢流检化验数据对应半小时前的生产数据；对于含空值、负值、零值等数据集，首先利用取空值算法去除数据集中所有含空值的样本，然后再采用相关离群点去除算法，去除数据集中所有含离群点的样本，目前有多种离群点的判定方法，本研究的离群值定义为与中位数相差超过 3 倍换算 S_{MAD}的元素，即超过 3 倍 S_{MAD}值的样本被认为是离群样本。

$$S_{\mathrm{MAD}}=\mathrm{c}\times\mathrm{median}(|\boldsymbol{A}-\mathrm{median}(\boldsymbol{A})|)$$

$$c=-1/\sqrt{2}\times\mathrm{erfcinv}(3/2) \tag{10}$$

式中：A 代表原始样本集；median( ) 函数是求解样本集中位数的函数；erfcinv( )是反误差函数。

图 7 和图 8 展示了加水量和溢流粒度部分数据集的预处理结果，溢流粒度是指磨矿细度为 0.075 mm的粒级占比。

最后，对采集工况对应的矿石可磨性检测数据的分布进行详细的分析，包括可磨性均值、方差、标准差和离散系数等，结果如表 1 所示。样本数据的分布合理性，对预测模型的评估影响很大。从表 1 数据可以看出，矿石可磨度均值在 13.377 4，离散系数为 0.147 9，采集样本离散度满足需求。

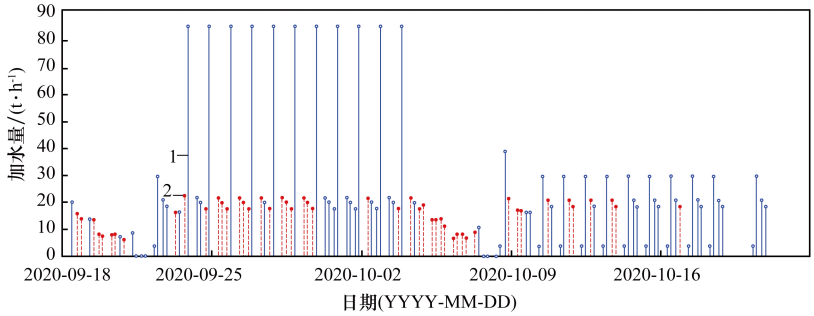
表 1　矿石可磨性样本数据分析

Table 1　Sample data analysis of ore grindability

均值	方差	标准差	离散系数
13.377 4	3.914 4	1.978 5	0.147 9

4.2　算法对比

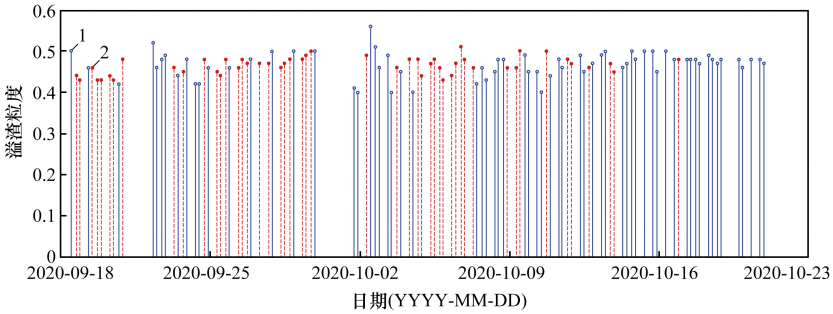
本文对比了 KPLSR、Gaussian_RBFNN 和 SAE 等算法在预测矿石可磨性方面的性能，性能指标主要包括均方根误差(S_{RMSE})、预测值与目标值的



1—离群点；2—预处理后。

图 7 预处理后的加水量数据

Fig. 7 Data of water after preprocessing



1—离群点；2—预处理后。

图 8 预处理后的溢流数据

Fig. 8 Overflow data after preprocessing

相关系数 R、决定系数 R^2、平均相对误差 S_{MRE}、相对误差率小于 5% 的命中率 $S_{Hitrate5}$ 和相对误差率小于 10% 的命中率 $S_{Hitrate10}$，具体如下。

4.2.1 KPLSR 的预测结果

KPLSR 采用 gauss 核函数，其中参数 σ 设置为 1，通过试凑法设置 KPLSR 的主成分对，得到 3 组结果较好的预测模型，如表 2 所示。由表 2 可知，当主成分对为 6 时，模型的预测性能最好，虽然 5% 误差的命中率稍低于 10 主成分对的模型，但预测值与目标值的相关性 R 接近 0.5。说明预测值的随机性小，跟随性好，因此这里认为当主成分对为 6 时，能获得最好的 KPLSR 模型来预测矿石的可磨性。

4.2.2 Gaussian_RBFNN 的预测结果

用基于高斯函数的 RBFNN 来预测矿石可磨度，只有在节点数为 2 时有测试结果，且 5% 偏差的命中率有 73.9%，其他组数的网络节点都遇到了奇异矩阵的问题，如表 3 所示。

表 2 KPLSR 的测试结果

Table 2 Test results of KPLSR

主成分对	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
10	0.633 920	0.365 816	−0.177 880	0.037 412	0.826 087	0.913 043
6	0.396 555	0.494 576	0.227 343	0.023 509	0.782 609	1.000
14	0.677 106	0.464 426	0.029 642	0.040 168	0.782 609	0.869 565

表 3 Gaussian_RBFNN 的测试结果

Table 3 Test results of Gaussian_RBFNN

节点	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
2	0.727 506	−0.109 63	−15.522 4	0.038 219	0.739 13	0.869 565

4.2.3　SAE 的预测结果

试验测试了 3 种结构，一层自编码 + FFNN，二层自编码 + FFNN，三层自编码 + FFNN，结果如下表 4 ~ 6 所示。

从表 4 ~ 6 可以看出，一层自编码 + FFNN 中 3∶2的结构获得了最佳预测结果。表 7 对比了 3 种算法的最佳结果。由表 7 可以看出，SAE 算法在应对复杂工业数据上更具有优势。

为了进一步验证 SAE 网络的性能，将数据集原来的 7∶3划分改为 6∶4和 8∶2，测试结果如表 8 所示。由表 8 可知，测试结果验证了 SAE 算法的稳定性。

表 4　SAE 的测试结果(一层自编码 + FFNN)

Table 4　Test results of SAE(one AE layer + FFNN)

网络结构	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
3∶6	0.726 397	0.217 935	0.019 021	0.044 240	0.695 652	0.913 043
3∶4	0.514 887	0.341 363	−0.154 180	0.028 164	0.869 565	0.956 522
3∶2	0.455 140	0.571 882	0.338 060	0.025 750	0.956 522	0.956 522
4∶2	0.655 341	0.253 780	0.056 937	0.037 885	0.826 087	0.913 043
2∶2	0.662 294	0.119 580	−0.103 200	0.037 279	0.826 087	0.913 043
4∶3	0.590 532	0.232 528	−0.090 660	0.034 136	0.826 087	1.000 000

表 5　SAE 的测试结果(二层自编码 + FFNN)

Table 5　Test results of SAE(two AE layers + FFNN)

网络结构	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
4∶2∶4	0.511 525	0.225 446	0.015 521	0.029 595	0.913 043	0.956 522
4∶2∶6	1.052 999	0.186 717	−3.171 850	0.070 964	0.304 348	0.739 130
4∶2∶2	0.637 230	0.272 679	−0.013 310	0.036 667	0.913 043	0.913 043
3∶2∶2	0.583 445	0.389 558	0.050 713	0.035 583	0.869 565	0.956 522
3∶2∶4	0.414 662	0.013 522	−0.325 960	0.026 150	0.913 043	1.000 000
3∶2∶6	0.829 167	0.331 275	0.002 994	0.046 881	0.695 652	0.913 043

表 6　SAE 的测试结果(三层自编码 + FFNN)

Table 6　Test results of SAE(three AE layers + FFNN)

网络结构	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
4∶3∶2∶2	0.750 714	0.235 506	−0.000 012	0.041 334	0.826 087	0.913 043

表 7　3 种算法最佳结果对比

Table 7　Comparison of the best results of three algorithms

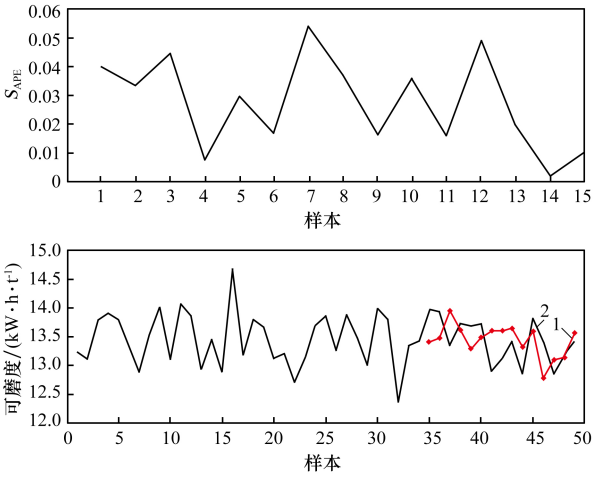
算法	网络结构	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
SAE	3∶2	0.455 140	0.571 882	0.338 060	0.025 750	0.956 522	0.956 522
KPLSR	6	0.396 555	0.494 576	0.227 343	0.023 509	0.782 609	1.000 000
Gaussian_RBFNN	2	0.727 506	−0.109 630	−15.522 400	0.038 219	0.739 130	0.869 565

表 8　SAE 预测结果(6∶4)

Table 8　The results of SAE(6∶4)

算法	网络结构	数据集划分	S_{RMSE}	R	R^2	S_{MRE}	$S_{Hitrate5}$	$S_{Hitrate10}$
SAE	3∶2	6∶4	0.685 836	0.445 150	0.316 640	0.042 760	0.645 161	0.935 484
SAE	3∶2	8∶2	0.508 877	0.713 905	0.372 984	0.028 650	0.941 176	1.000 000

综上，将 SAE 最佳训练模型应用于铁矿石可磨度的智能识别。根据选定的 30% 的测试数据，预测的矿石可磨度与实际检测值的平均绝对误差百分比仅为 2.51%，偏差 5% 以内的预测命中率可达 93.33%。具体展示结果见图 9。



1—预测值；2—检测值。

图 9 可磨度测试集的识别效果

Fig. 9 Recognition effect of grindability test sets

5 矿石可磨度智能识别系统和应用效果

5.1 矿石可磨度智能识别系统

在线识别主要实现生产数据的动态展示和相应矿石可磨度的预测，其展示效果如图 10 所示。

该区域上方为加水量、给矿量、磨机电流、磨机功率的历史数据折线图，最大显示数据量为 800 条，该区域内可以使用鼠标进行拖动、缩放等查看操作，鼠标悬停可以查看该时间点的具体数值。该区域下方为当前时刻的加水量、给矿量、磨机电流、磨机功率、溢流粒度、钢球单耗、可磨度预测的数值，以及当前数据的时间。

5.2 工业化应用效果

对系统应用生产线上入磨机前矿石进行了不定期采样，在实验室测定其 Bond 功指数，并与系统智能识别结果进行对比，对比结果如表 9 所示。由表 9 可知，与 30 批次的生产线上矿石可磨度测定结果相比，智能识别系统给出的矿石可磨度绝对值误差平均为 0.61 kW·h/t，平均误差率约为 4.79%，基本满足矿石可磨度智能识别的预期目标。

表 9 智能识别系统结果分析

Table 9 Analysis results of intelligent recognition system

序号	可磨度（化验值）	可磨度（预测值）	绝对误差值	绝对误差率/%
1	13.42	13.59	0.174 4	1.30
2	13.15	12.86	0.289 1	2.20
3	12.96	13.06	0.095 4	0.74
4	12.09	13.38	1.289 1	10.66
5	12.68	13.64	0.956 3	7.54
6	13.29	12.81	0.480 4	3.61
7	12.97	12.63	0.342 5	2.64
8	12.02	12.45	0.425 1	3.54
9	12.75	12.81	0.064 6	0.51

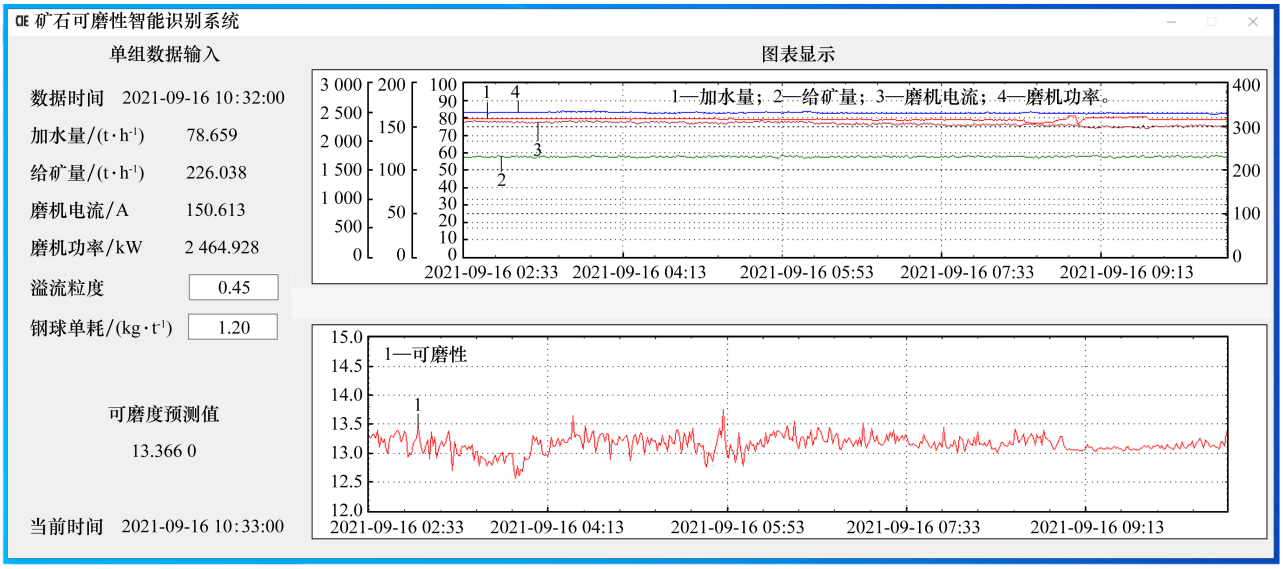


图 10 矿石可磨度智能识别系统界面

Fig. 10 System interface of ore grindability intelligent identification

表 9（续）

序号	可磨度（化验值）	可磨度（预测值）	绝对误差值	绝对误差率/%
10	11.85	12.51	0.657 3	5.55
11	11.69	13.16	1.467 3	12.55
12	11.69	13.33	1.640 9	14.04
13	13.51	13.35	0.163 1	1.21
14	14.00	13.16	0.838 2	5.99
15	13.12	12.72	0.401 9	3.06
16	13.60	13.31	0.289 1	2.13
17	12.91	14.22	1.312 1	10.16
18	12.67	13.57	0.897 8	7.09
19	12.80	14.04	1.239 4	9.68
20	12.70	13.88	1.180 9	9.30
21	14.01	13.77	0.244 3	1.74
22	13.14	13.27	0.130 2	0.99
23	13.47	13.27	0.198 8	1.48
24	13.61	13.26	0.346 1	2.54
25	12.64	12.91	0.266 6	2.11
26	13.01	13.52	0.5128	3.94
27	13.59	13.75	0.163 7	1.20
28	13.79	13.45	0.335 9	2.44
29	13.34	12.56	0.776 6	5.82
30	13.66	12.56	1.096 6	8.03
平均	—	—	0.609 2	4.79

6　结　论

（1）通过对磨矿过程的详细分析，得出磨矿过程运行指标的变化能从一定程度上反应矿石可磨度变化的结论，但磨矿过程是一个循环连续的过程，引起运行指标变化的因素很多，各因素间相互耦合严重，很难从机理的角度理顺各参数的变化与矿石可磨度之间的因果关系。

（2）提出了以数据驱动的矿石可磨性智能识别方法的整体结构，并着重分析了 KPLSR 和 SAE 两种可提取关键特征的算法机理。以实际采集的磨矿过程数据为样本，经过预处理后用于多种智能算法的训练测试，通过每个算法的多组试验得到该算法的最佳模型，然后再进行算法间的比较，算法性能采用多指标评估，最终 SAE 算法在矿石可磨性的预测中表现出最好的性能。

（3）矿石可磨性智能识别系统选择测试试验中性能最佳的算法结构作为在线识别模型，在工业化应用中，达到了矿石可磨度智能识别的需求。

（4）生产数据驱动的矿石可磨度智能识别技术，作为软测量预测识别工业化应用的一个成功案例，其共性技术可以拓展到其他应用场景，解决同类型参数长期无法在线获取的难题。

参考文献：

[1] BOND F C, WANG J T. A New theory of comminution [J]. Mining Engineering, 1950, 62(8): 871 – 878.

[2] ROWLAND C A. Grinding calculations related to the applications of large rod and ball mills abstract[J]. Canada Mining Jounal, 1972, 93(6): 48 – 53.

[3] 熊维平，翁伟雄，周忠尚．影响邦德功指数测定精度因素的研究[J]．金属矿山，1983(10)：35 – 37.
XIONG Weiping, WENG Weixiong, ZHOU Zhongshang. Study on the factors affecting the determination accuracy of Bond's work index[J]. Metal Mine, 1983(10): 35 – 37.

[4] BERRY T F, BRUCE R W. A simple method of determing the grindability of oress[J]. Canada Mining Journal, 1966, 87(7): 63 – 65.

[5] KARRA V K. Simulation of the Bond grindability test[J]. CIM Bulletin, 1981, 102(3): 195 – 199.

[6] KAPUR P C. Analysis of the Bond grindability test[J]. Transactions of the institution of mining and metal lurgy, 1970, 79: 103 – 108.

[7] 赖复兴．邦得可磨性试验模拟功指数的简化算法[J]．有色金属(选矿部分)，1982(3)：44 – 48.
LAI Fuxing. Simplified algorithm for simulation work index of Bond grindability test[J]. Nonferrous Metals (Mineral Processing), 1982(3): 44 – 48.

[8] 熊维平，翁伟雄，周忠尚．邦德功指数测定的计算机仿真方法[J]．有色金属，1984(4)：28 – 34.
XIONG Weiping, WENG Weixiong, ZHOU Zhongshang. A computer simulation method for the determination of the bond work index[J]. Nonferrous Metals, 1984(4): 28 – 34.

[9] 陈炳辰，周凌嘉，仲崇波．邦德功指数模拟计算方法的研究[J]．金属矿山，1990(8)：36 – 39.
CHEN Bingchen, ZHOU Lingjia, ZHONG Chongbo. Research on simulation calculation method of Bond's

work index[J]. Metal Mine,1990(8):36 – 39.

[10] 仲崇波,王成功,陈炳辰. 邦德棒磨功指数模拟计算方法的研究[J]. 中国矿业,1997,6(2):48 – 52.
ZHONG Chongbo, WANG Chenggong, CHEN Bingchen. A study of simulating calculation method of the rod-mill's bond work index[J]. China Mining Magazine, 1997,6(2):48 – 52.

[11] 吴建明. Bond 粉磨功指数研究与应用的进展[J]. 有色设备,2005,19(3):1 – 3.
WU Janming Progress in research and application of bond grinding work index[J]. Non-ferrous Metallurgical Equipment,2005,19(3):1 – 3.

[12] KADLEC P, GABRYS B, STRANDT S. Data-driven Soft Sensors in the process industry[J]. Computers & Chemical Engineering,2009,33(4):795 – 814.

[13] 王晓. 基于深度学习的复杂工业软测量理论与方法[D]. 西安:西安理工大学,2020.
WANG Xiao. Theory and method of complex industrial soft sensing based on deep learning[D]. Xi'an: Xi'an University of Technology,2020.

[14] YUAN X F, HUANG B, WANG Y L, et al. Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE [J]. IEEE Transactions on Industrial Informatics,2018, 14(7):3235 – 3243.

[15] YUAN X F, ZHOU J, HUANG B, et al. Hierarchical quality-relevant feature representation for soft sensor modeling: A novel deep learning strategy[J]. IEEE Transactions on Industrial Informatics, 2020, 16 (6): 3721 – 3730.

[16] MA M, KHATIBISEPEHR S, HUANG B. A Bayesian framework for real-time identification of locally weighted partial least squares[J]. AIChE Journal,2015,61(2): 518 – 529.

[17] FORTUNA L, GIANNONE P, GRAZIANI S, et al. Virtual instruments based on stacked neural networks to improve product quality monitoring in a refinery[J]. IEEE Transactions on Instrumentation and Measurement,2007, 56(1):95 – 101.

[18] KANEKO H, FUNATSU K. Application of online support vector regression for soft sensors[J]. AIChE Journal, 2014,60(2):600 – 612.

[19] YU J, HONG C Q, RUI Y, et al. Multitask autoencoder model for recovering human poses[J]. IEEE Transactions on Industrial Electronics,2018,65(6):5060 – 5068.

广 告 目 次

封面	湖南中冶长天重工科技有限公司	前彩 20 – 21	北京顺泰克科技有限公司
封二	襄阳中和机电技术有限公司	前彩 22 – 23	河北华通重工机械制造有限公司
封三	《烧结球团》杂志征订启事	前彩 24	焦作市迈科冶金机械有限公司
封底	河北同业冶金科技有限责任公司	前彩 25	上海铭伦冶金科技开发有限公司
前彩 1	中冶长天炉窑工程技术公司	前彩 26 – 27	唐山市磁石矿冶科贸有限公司
前彩 2 – 3	鞍山市昱昊机械设备有限公司	前彩 28 – 29	北京金都泰拓冶金技术研究有限公司
前彩 4 – 5	湖南德尚源耐磨工业有限公司	前彩 30	SEW – 传动设备(天津)有限公司
前彩 6 – 7	湖南晟清环保科技有限公司	前彩 31	江阴市创裕机械有限公司
前彩 8	江苏中鼎环境工程股份有限公司	前彩 32	成都利君实业股份有限公司
前彩 9	罗迪格(北京)机械设备有限公司	前彩 33	长沙天为华信环保装备有限公司
前彩 10 – 11	中冶长天(长沙)智能科技有限公司	前彩 34 – 35	爱立许集团(中国)有限公司
前彩 12 – 13	朝阳重型机器有限公司	前彩 36	青岛百瑞福橡塑有限公司
前彩 14 – 15	北京艾瑞机械有限公司	后彩 1	湖南精城特种陶瓷有限公司
前彩 16	南京三埃工控有限公司	后彩 2	鞍山市明阳金属材料有限公司
前彩 17	平罗县国宁活性炭有限公司	后彩 3	湖北中船气体有限公司
前彩 18 – 19	湖南中冶长天重工科技有限公司	后彩 4	中冶京诚工程技术有限公司